# Fixed-Effect Estimation of Highly-Mobile Production Technologies

**William C. Horrace***
Syracuse University

**Kurt E. Schnier**
Georgia State University

September, 2009

**Key Words:** Panel data, time-varying individual effect, spatial econometrics, fisheries, agriculture, heteroskedasticity.

**JEL Codes:** C23, D24, N50

* Corresponding Author. Syracuse University, 426 Eggers Hall, Syracuse NY, 13244; phone: 315-443-9061; fax: 315-443-1081; email: whorrace@maxwell.syr.edu. Thanks to Antonio Alvarez, Badi Baltagi, David Castilla, Stacey Chen, Jaesung Cho, Chihwa Kao, Bill Greene, Terry Kinal, Nicolai Kuminoff, Kajal Lahiri, George Monokroussos, Chris Parmeter, Ozgen Sayginsoy, and Loren Tauer for comments. Thanks to Michael Eriksen for the GIS plots. Errors are ours.

## 0. Abstract

We consider fixed-effect estimation of a production function where inputs and outputs vary over time, space, and cross-sectional unit. We exploit variability in the spatial dimension to identify time-varying individual effects, without parametric assumptions on the effects. Under a weak exogeneity assumption estimates of marginal products and individual effects are unbiased. We also discuss sufficient conditions for consistency and asymptotic normality along the spatial dimension. We apply our results to a production function of bottom trawl fishing vessels in the flatfish fisheries of the Bering Sea. We find significant spatial variability of output (catch) which we exploit in estimation of a harvesting function. We conclude that vessel individual effects did not change across the period 2002 to 2004. We apply the theory of ranking and selection to determine that individual effects are not statistically significant across vessels.

## 1. Introduction

Consider the econometric fixed-effect model:

$$y_{it} = \alpha_{it} + x_{it}\beta + z_i\gamma + v_{it}\,, \qquad i = 1,\ldots,N\,,\ t = 1,\ldots,T\,,$$

where $i$ indexes individual or cross-sectional unit, and $t$ indexes time. Notice that the individual effects, $\alpha_{it}$, vary over time. The earliest specifications of this model were identified by the restriction $\alpha_{it} = \alpha_i$ for all $t$, producing the common panel data specification (see Mundlak, 1978; MaCurdy, 1981; and Chamberlain, 1984). To relax this restriction a series of papers parameterize the time-varying effects into an individual component and a time component, so that the temporal pattern is fixed across individuals or groups of individuals. See Cornwell, Schmidt, and Sickles (1990), Kumbhakar (1990), Battese and Coelli (1992), Lee and Schmidt (1993), Cuesta (2000), Ahn, Lee, and Schmidt (2001), Han, Orea and Schmidt (2005), and Lee (2005).

An excellent discussion of time-varying individual effects models, their underpinnings, estimation, and applicability is provided in the introduction of Ahn, Lee, and Schmidt (2001). In particular they relate these models to the work of Kiefer (1980), Holtz-Eakin et al. (1988), and Chamberlain (1992). They also discuss their application to rational expectations models (Hall and Mishkin, 1982; Shapiro, 1984; and Keane and Runkle, 1992), production function estimation (Schmidt and Sickles, 1984; and Lee and Schmidt, 1993), and estimation of earnings equations where unobserved ability might vary with time due to a time-varying implicit price of ability. There is also a sizeable Bayesian literature that addresses panel data estimation of panel data models and production functions. However, Bayesian approaches are not directly comparable to the frequentist approaches considered herein, so while the Bayesian literature is certainly important, it will not be discussed here.[1]

The intent of this research is to relax the parametric assumptions on time-varying individual effects, and exploit spatial variation of economic agents to identify and estimate the model with a 'within' transformation and ordinary least-squares. Our primary interest is production function estimation, but our results could also be applied in any of the aforementioned empirical settings, as long as agents are highly-mobile, location-specific data are observed, and the variability of output is statistically relevant along the spatial dimension.

While most production technologies are fixed (in the short-run), one can envision technologies that are not. The example we discuss in detail is the fishery, where fishing vessels harvest fish in different spatial locations of the sea and where spatial variability of harvest is statistically meaningful. Other examples of highly-mobile technologies are: police cruisers arresting criminals in different locations of a city, taxis competing for fares, sales forces

[1] For Bayesian treatments of panel data frontier models see, for example, Fernandez et al. (2002), Tsionas (2002), Kim and Schmidt (2000), and Koop et al. (1997).

mobilized to serve clients, farm combining operations that move from south to north over the course of a growing season, or natural gas and oil drilling operations.[2] Here, the dependent variable (production) may be observed over time, space, and individual (i.e., $y_{its}$). With adequate spatial variability in the factors of production ($x_{its}$) the time-varying individual effects ($\alpha_{it}$) can be modeled without parameterization. In fact, $\beta$ in the linear model,

$$y_{its} = \alpha_{it} + x_{its}\beta + z_{it}\gamma + w_i\delta + v_{its},$$

can be estimated with a simple 'within' transformation, where within-cell averages are taken over the spatial dimension $s$ (i.e., $y_{its} - \bar{y}_{it}$). Under a weak exogeneity assumption on the errors and regressors, the within estimate is unbiased. In this paper, we consider only 'within' estimation and deal with several perplexing issues related to it. The most difficult of which is that the parameters of space invariant production factors, $z_{it}$ and $w_i$, are not identified. (This is particularly vexing when estimates of the individual effects are desired. See Feng and Horrace, 2007.) This problem is tackled by recognizing that mobile technologies are usually engaged in the harvesting of some natural resource or moving to where the stock of raw materials of production are most abundant (e.g., fishing vessels harvest fish, police 'harvest' criminals, and taxis 'harvest' fares). If the resource stocks (fish, criminals, etc.) are observable within each spatial location and vary over space, then we posit a *harvesting function*, in the spirit of Schaefer (1957), which *interacts* space-varying stock with the factors of production. As such, all the factors of production are (effectively) space-varying and are, thus, identified. Identification hinges critically on the fact that individual effects do not vary over space (i.e., $\alpha_{it}$ remains fixed across $s$). Identification also hinges on the assumption that resource stocks are exogenous,

[2] Frequent relocation of capital to maximize profits (or minimize cost) is an inevitability as the time dimension of a panel become large (in the long-run). Consider the flow of capital from the northern U.S. to the southern U.S. over the last twenty years. Of course, large *T* presents many challenges not addressed in this research, as we fix *T*.

which we assume throughout this paper. Of course, if stocks are endogenous then some form of instrumental variables estimation is needed.[3] For our example, the measure of resource stock is, indeed, exogenous. These complications are discussed in the sequel.

Finally, we consider asymptotics in the spatial dimension and show that our within estimate of $\beta$ is a linear case of Pinkse, Slade and Shen (2006) and Pinkse, Shen and Slade (2007), so consistency and asymptotic normality hold for our estimates under their spatial mixing assumptions. Both of the aforementioned papers are conceptualized for asymptotics along the cross-sectional dimension, where each agent endogenously selects a spatial location, and the number of agents and, hence, spatial locations grow. Our concept of a fixed number of agents moving over a growing number of spatial locations is slightly different, but their ideas still apply. The primary difference between our work and theirs is our indexing strategy, that allows us to specifically model time-varying heterogeneity. This could only be accomplished in a meaningful way using our idea of a fixed number of agents moving over ever-growing area. There are, however, certain practical drawbacks to this concept, and we discuss them in the sequel.

It should be noted that three-dimensional panels have been considered in empirical work in the past, but our model is unique in two ways. First, ours is the first to consider "within" estimation. Other papers have considered the " least squares dummy variable estimator" (LSDV) and ignore certain econometric nuances that we discuss in detail. For example, Parsley and Wei (2005) consider a LSDV regression of the variability of prices on traded goods ($i$), over time ($t$), and across cities ($s$) in the U.S. and Japan. Second, our three dimensions: cross-section, time, and space, are uniquely distinct features of the data. Other papers have added a third dimension to a panel that is not distinct from the others. For example, Davies and Lahiri (1995) consider a

[3] In some fisheries harvesting may serially deplete resource stocks, making resource stock endogenous. This is not a serious concern in the fishery we consider in the sequel, however serial depletion may be present in some of the aforementioned examples, such as policing where intensive criminal 'harvesting' may deplete resource stocks.

panel of forecasters (*i*), in time period (*t*), for forecast horizons (*h*). However, their third dimension (*h*) is merely a subdivision of time (*t*). Eilat and Einav (2004) develop a three-dimensional model of tourism flows over country of origin (*o*), destination country (*d*) and time (*t*). However, again the dimensions *o* and *d* are not uniquely distinct. When subdivisions of geographical entities are considered, the subdivisions are necessarily static. For example, there are many papers (e.g., Valletta, 1993 or Fleck, 1999) that analyze data over U.S. cities or counties, within states, over time. These are usually treated as two-dimensional panels (cities or counties over time) with state dummies, because there is no sense in which cities or counties can move across states in the way that they 'move' across time. Our data are unique in that each cross sectional unit can move across both time and space.[4]

Most spatial econometric innovations in the last ten years are conceptualized for fixed (or nearly-fixed) economics agents. This is not entirely unrealistic since in the short-run economic agents and capital remain in a fixed location. For example, Conley's series of spatial econometric papers are all based on a one-shot view of space, where agents are not changing position. See Conley (1999), Conley and Dupor (2003), Conley and Ligon (2002), and Conley and Topa (2002). Also, papers based on fixed weighting matrices do the same. For example, see Kelijian and Prucha (1999 and 2001). In these papers, the presumption is that there is not enough mobility over time for space to be considered as another source of variability in the data. Indeed, we contend that they are either assuming that resources are fixed (e.g., immobile capital or natural resource), or that the time dimension is not large enough for mobility to be considered a reasonable assumption. Therefore, by relaxing the assumptions of spatially fixed inputs, our model makes a unique contribution to the literature on spatial econometrics.

---

[4] This is also a unique feature of Parsley and Wei (2001).

The paper is organized as follows. The next section defines the harvesting function, and discusses identification, estimation and unbiasedness. Section 3 discusses practical issues concerning asymptotics and aggregation and develops the assumptions necessary for asymptotic normality, based on the central limit theorem of Pinkse, Shen and Slade (2007). In section 4, we present an example: estimation of a production function of bottom trawl fishing vessels in the flatfish fisheries of the Bering Sea. The last section concludes and makes suggestions for future research.

## 2. Specification and Algebra

In what follows, we couch the discussion in terms of the example of interest, Bearing Sea flatfish fisheries. However, the discussion is relevant to all the aforementioned highly-mobile technologies. Define the Cobb-Douglas harvesting function:

$$y_{its} = A_{it}\{x_{its}^{\beta} z_{it}^{\gamma} w_i^{\delta}\}^{b_{ts}} \exp(v_{its}) \quad i = 1,\ldots,N\ ,\ t = 1,\ldots,T\ ,\quad s = 1,\ldots,S_{it}\ ,$$

where *s* indexes spatial location fished, $i$ indexes the vessel, and *t* indexes time. Notice that we allow the number of spatial locations, $S_{it}$, to vary over $i$ and $t$; this is the spatial equivalent of an unbalanced panel. We make explicit the fact that the exogenous inputs to the harvesting function may be space-invariant ($z_{it}$), or possibly space- and time-invariant ($w_i$). The $b_{ts}$ is an *observed* time- and space-varying exogenous factor of harvesting, which doesn't vary over $i$ and is a limiting factor for all harvesting inputs. In our fisheries context this would be the fish density (biomass) in a given location and time period. The idea is that fishing stocks are exogenous (as we shall see), and production efforts are only successful when fish are present. The exogeneity of $b_{ts}$ may be called into question for many applications. In this context we think of endogeneity as coming from the decision of 'where to harvest.' That is, the location of the means of production is a key choice variable in the optimization problem. For example, cabbies elect to search for fares where population density is highest, and police forces patrol more in areas where the crime rate is highest, so production (output) effects the location decision, which is correlated to stocks of harvestable resources in each location.[5]

Notice that the inputs to fishing are affected by the biomass through the exponent $b_{ts}$ and that technical change, $A_{it}$, is constant over all spatial locations and is, consequently, unaffected by the biomass in the spatial location (it is not raised to the $b_{ts}$ power) . This is critical to identification for 'within' estimation of the model.[6] Taking logs yields the following log-transformed production function:

$$\ln y_{its} = \ln A_{it} + b_{ts} \ln x_{its}\beta + b_{ts} \ln z_{it}\gamma + b_{ts} \ln w_i \delta + v_{its} .$$

Let $\alpha_{it} = \ln A_{it}$, $Y_{its} = \ln y_{its}$, $X_{its} = b_{ts} \ln x_{its}$, $Z_{its} = b_{ts} \ln z_{it}$, and $W_{its} = b_{ts} \ln w_i$, then:

$$Y_{its} = \alpha_{it} + X_{its}\beta + Z_{its}\gamma + W_{its}\delta + v_{its}, \quad i = 1,...,N, \ t = 1,...,T, \quad s = 1,...,S_{it} . \qquad (1)$$

This is just a fixed-effect specification, but the beauty of it is that ALL the regressors vary over *s* (due to their interactions with $b_{ts}$, which *does* vary over *s*). Therefore, all the parameters ($\beta$, $\gamma$, and, $\delta$) are identified by 'within' estimation. The point is that inputs alone do not catch fish; it is the interaction of the biomass or density of fish with the production inputs that catch fish. As such, inputs that do not vary with spatial location (like a vessel’s size) can be interacted with biomass in different locations to identify the parameters of the model. This is similar in spirit to Wooldridge's 'solution' to time invariant regressors in the usual fixed-effect model: they are not allowed "unless they are interacted with time varying variables, such as time dummies" (Wooldridge, 2002, p269). Although, in this case the interactions are well-justified, as it would seem that the marginal products of fishing inputs would equal zero when there were no fish to

[5] In particular we do not view the endogeneity as coming directly from the harvesting. That is aggressive harvesting does not lower the fish stocks in any appreciable way in the short-run.
[6] It is not critical if we assume a parametric form for ln$A_{it}$ and perform GMM.

catch but would be very large when there are many fish to catch (particularly when they are being caught in trawling nets). Consequently, interaction of inputs with biomass makes sense both empirically and theoretically.[7]

One could also envision a specification where biomass (alone) enters the harvesting function log-linearly and is multiplied by a marginal product parameter for estimation. This presents no additional problems in the estimation. However, the specification would imply the Cobb-Douglas harvesting functions:

$$y_{its} = A_{it}\{x_{its}^{\beta} z_{it}^{\gamma} w_i^{\delta} b_{ts}\}^{b_{ts}} \exp(v_{its}) \text{ or } y_{its} = A_{it} b_{ts}\{x_{its}^{\beta} z_{it}^{\gamma} w_i^{\delta}\}^{b_{ts}} \exp(v_{its}),$$

which seems somewhat redundant because of $b_{ts}$ occurring twice in the form. These functions are within the realm of possibilities, but are not considered in what follows. It should also be noted that the Cobb-Douglas harvesting function is easily generalized to a trans-log specification, with variable interactions across all three dimensions in the spatial panel.

Consider the specification in equation 1 in more detail. We have implicitly assumed that the inputs ( $X_{its}$, $Z_{its}$, and $W_{its}$ ) and the parameters ( $\beta$, $\gamma$, and, $\delta$ ) are scalars. Let's make things more general. First, let $Y_{its}$ and $v_{its}$ be scalars. Let $X_{its}$, $Z_{its}$, and $W_{its}$ be $(1\times k)$, $(1\times g)$, and $(1\times d)$ row vectors, respectively. Let $\beta$, $\gamma$, and $\delta$ be $(k\times 1)$, $(g\times 1)$, and $(d\times 1)$ column vectors, respectively. Let,

$$\underset{(1\times[k+g+d])}{\tilde{X}_{its}} = \begin{bmatrix} X_{its} & Z_{its} & W_{its} \end{bmatrix} \text{ and } \underset{([k+g+d]\times 1)}{\beta_*'} = \begin{bmatrix} \beta' & \gamma' & \delta' \end{bmatrix},$$

Then, our equation becomes

$$Y_{its} = \alpha_{it} + \tilde{X}_{its}\beta_* + v_{its}.$$

Defining the variables demeaned over the spatial dimension,

---

[7] We could also follow Wooldridge and interact all variables with location dummies. However, we desire asymptotics along the spatial dimension, so spatial dummies are infeasible due to incidental parameters.

$$Y^{+}_{its} = Y_{its} - S^{-1}_{it}\sum_{s=1}^{S_{it}} Y_{its}\ ,\ \ \tilde{X}^{+}_{its} = \tilde{X}_{its} - S^{-1}_{it}\sum_{s=1}^{S_{it}} \tilde{X}_{its}\ \text{, and }\ v^{+}_{its} = v_{its} - S^{-1}_{it}\sum_{s=1}^{S_{it}} v_{its}\ ,$$

our demeaned equation is,

$$Y^{+}_{its} = \tilde{X}^{+}_{its}\beta_{*} + v^{+}_{its}\ ,\ i = 1,...,N\ ,\ t = 1,...,T\ ,\ s = 1,...,S_{it}\ . \qquad (2)$$

Under appropriate exogeneity assumptions and regularity conditions on the regressors (e.g., see White, 1984) ordinary least-squares (OLS) of this equation produces unbiased estimate,

$$\hat{\beta}_{*} = \left(\sum_{i}^{N}\sum_{t}^{T}\sum_{s}^{S_{it}} \tilde{X}^{+'}_{its}\tilde{X}^{+}_{its}\right)^{-1}\left(\sum_{i}^{N}\sum_{t}^{T}\sum_{s}^{S_{it}} \tilde{X}^{+'}_{its}Y^{+}_{its}\right).$$

Notice that all elements of $\hat{\beta}_{*}$ are identified, because all elements of $\tilde{X}^{+}_{its}$ are space-varying through interactions with biomass, $b_{ts}$. Let the 'within' residual be $\hat{v}^{+}_{its} = Y^{+}_{its} - \tilde{X}^{+}_{its}\hat{\beta}_{*}$. Then an unbiased estimate of $\alpha_{it}$ is,

$$\hat{\alpha}_{it} = S^{-1}_{it}\sum_{s=1}^{S_{it}} \hat{v}^{+}_{its}\ .$$

The usual panel data version of this model ($S_{it} = 1$) is commonly employed to estimate time-invariant technical efficiency from a stochastic frontier model (see Schmidt and Sickles, 1984). In our generalized panel data case ($S_{it} > 1$) we can estimate time-varying technical efficiency. Let $\alpha_{it} = \eta - u_{it}$, where $\eta$ is an "overall" fixed intercept parameter, and $u_{it}$ is a non-negative parameter, representing time-varying technical inefficiency. Then, following Schmidt and Sickles (1984), relative time-varying inefficiency is $u_{it} = \max_{j,w}\alpha_{jw} - \alpha_{it}$, and can be estimated as $\hat{u}_{it} = \max_{j,w}\hat{\alpha}_{jw} - \hat{\alpha}_{it}$. A consistent normalization of technical efficiency is $T\hat{E}_{it} = \exp\{-\hat{u}_{it}\} \in (0,1]$. An alternative measure is $u^{*}_{it} = \max_{j}\alpha_{jt} - \alpha_{it}$, estimated as

$\hat{u}_{it}^{*} = \max_{j} \hat{\alpha}_{jt} - \hat{\alpha}_{it}$, which implies a relatively inefficient $i$ within each period $t$ and which yields technical efficiency estimate $T\hat{E}_{it}^{*} = \exp\{-\hat{u}_{it}^{*}\}$.

**3. Spatial Asymptotics**

Under a weak exogeneity assumption on the regressors, the marginal effects and the individual effects estimates of equation 1 are unbiased, so our discussion of asymptotics is intended to facilitate inference when the errors are non-normally distributed or when robust inference is necessary. The latter situation may arise when the data are aggregated. Aggregation may be necessary for data from highly-mobile technologies, as we shall see. In what follows we let $S_{it} = S$ for all $i, t$ without loss of generality. It should be noted that if inference on $\beta_{*}$ is the focus, then asymptotics can be conceptualized along the *N* or *T* dimensions or fixed *S*, and what follows may be unnecessary. For inference on $\alpha_{it}$, some notion of asymptotics along the spatial dimension is necessary, and we discuss this next.

If we think of physical space (say, the sea) as a two-dimensional rectangular integer lattice, then production can move to any of $S$ spatial regions within a given time period, $t$. Given this, we can think of asymptotics in two extreme ways: either a) the surface area of the lattice (domain) expands and the area of the individual locations is fixed as $S \to \infty$, or b) the area of the lattice (domain) is fixed, and the number of spatial locations increases while their area size decreases, as $S \to \infty$. Following Cressie (1993) we call the former "increasing-domain asymptotics" and the latter "infill asymptotics."[8]

There are spatial central limit theorems in existence. In the economics literature the most recent results are in Pinkse, Shen and Slade (2007), who also apply these results to inference in a

[8] Infill asymptotics can be motivated by recent advancements in the resource economics literature which divide a fishery into spatially distinct “patches” (Sanchirico and Wilen, 1999, 2005). Each patch is defined by the ecological characteristics of the resource and the degree of resource heterogeneity present.

non-linear regression model in Pinkse, Slade and Shen (2006). Their CLT results are based on the Bernstein (1927) grouping strategy, whereby spatial locations are partitioned in such a way that spatial dependence decreases as the number of locations increases. "Such partitioning does not actually have to take place; there merely must be the possibility to do so" (Pinkse, Shen and Slade, forthcoming p. 219). For completeness we now present assumptions for our asymptotics in the model of equation 1. In what follows all our arguments are based on $S \to \infty$.

Let $\tilde{X}_s^+$ be the matrix of observations of $\tilde{X}_{its}^+$ for all $i, t$ in location $s$, so it is a $N \times T$ matrix. Similarly for $v_s^+$ with covariance matrix $\Omega$. Now assume the following.

A1. $E(\tilde{X}_s^{+'} v_s^+) = 0$.

A2. $\Gamma = V\left( S^{-1/2} \sum_{s=1}^{S} \tilde{X}_s^{+'} v_s^+ \right) = S^{-1} \sum_{s=1}^{S} \tilde{X}_s^{+'} \Omega \tilde{X}_s^+$ is $O_p(1)$ and uniformly positive definite.

A3. For any vector $\lambda$ with $\lambda\lambda' = 1$, assumptions A-C of Pinkse , Shen and Slade (2007) are satisfied for:

$$\lambda \left( \sum_{s=1}^{S} \tilde{X}_s^{+'} \Omega \tilde{X}_s^+ \right)^{-1/2} \left( \sum_{s=1}^{S} \tilde{X}_s^{+'} v_s^+ \right).$$

Letting the $k^{\text{th}}$ element of vector $\tilde{X}_{its}^+$ be the scalar $\tilde{X}_{its;k}^+$,

A4. $\left| Cov(\tilde{X}_{its;k}^+ \tilde{X}_{i^*t^*s;k^*}^+ , \tilde{X}_{jmr;g}^+ \tilde{X}_{j^*m^*r;g^*}^+) \right| \le \rho_{sr} \sqrt{V(\tilde{X}_{its;k}^+ \tilde{X}_{i^*t^*s;k^*}^+)} \sqrt{V(\tilde{X}_{jmr;g}^+ \tilde{X}_{j^*m^*r;g^*}^+)}$ where

$\mathrm{P} = \max_{s \le S} \sum_{r=1}^{S} \rho_{sr}$ and $\lim_{S \to \infty} \frac{\mathrm{P}}{S} = 0$, for all $i$, $j$, $t$, $m$, $k$, $g$, $i^*$, $j^*$, $t^*$, $m^*$, $k^*$, $g^*$, $s$, $r$.

A5. $V\left( S^{-1} \sum_{s=1}^{S} \tilde{X}_s^{+'} \tilde{X}_s^+ \right)$ is $O_p(1)$ and uniformly positive definite.

A6. $M = E\left( S^{-1} \sum_{s=1}^{S} \tilde{X}_s^{+'} \tilde{X}_s^+ \right)$ is $O_p(1)$ and uniformly positive definite.

Assumption A1 is a standard exogeneity assumption. A sufficient condition for A1 is $E(v_s \mid \tilde{X}_s) = 0$. Assumption A2 allows for arbitrary covariance structure for the error (this is important not only for spatial dependence but for aggregation issues that we discuss in the sequel). The conditions necessary for A3 to hold are those that ensure spatial 'mixing' for sums of the $\tilde{X}_s^{+'} v_s^{+}$. The most important condition is a covariance-variance inequality with mixing constants similar to A4, but for the elements of the product of $\tilde{X}_{its}^{+}$ and $v_{its}^{+}$.

The spatial mixing condition in A3 is based on a blocking strategy (Bernstein, 1927), which partitions the sample into groups (blocks) and subgroups (sub-blocks), such that as the sample size grows, dependence between subgroups within a group becomes negligible. First, it is important to note that the blocking strategy is not limited to partitioning physical space; it is more general. Indeed, Pinkse, Shan and Slade (2007) state that the blocks "do not have to be blocks; they are judiciously chosen subsets of observations." Second, the blocking "does not actually have to take place; there merely must be the possibility to do so." For our application we think of the blocks as a partitioning of the sea and the asymptotics are for an increasing domain.

Assumptions A1, A2 and A3 ensure that, $S^{-1/2} \sum_s \tilde{X}_s^{+'} v_s^{+} \xrightarrow{d} N(0, \Gamma)$. Assumption A4 is the linear version of equation 25 of Pinkse, Slade and Shen (2006) and is a spatial mixing condition. The constants $\rho_{sr}$ limit the amount of spatial dependence to ensure convergence of the second moment matrix. A5 and A6 are standard. Together A4, A5 and A6 ensure that $S^{-1} \sum_2 \tilde{X}_s^{+'} \tilde{X}_s^{+} - M \xrightarrow{p} 0$. To see this, notice that A4 implies:

$$S^{-1}\sum_s E\left[\tilde{X}^+_{its;k}\tilde{X}^+_{i^*t^*s;k^*} - E\left(\tilde{X}^+_{its;k}\tilde{X}^+_{i^*t^*s;k^*}\right)\right]^2$$

$$\leq S^{-2}\sum_{s,r}\left|Cov(\tilde{X}^+_{its;k}\tilde{X}^+_{i^*t^*s;k^*},\tilde{X}^+_{jmr;g}\tilde{X}^+_{j^*m^*r;g^*})\right|$$

$$\leq S^{-2}\rho_{sr}\sqrt{V(\tilde{X}^+_{its;k}\tilde{X}^+_{i^*t^*s;k^*})}\sqrt{V(\tilde{X}^+_{jmr;g}\tilde{X}^+_{j^*m^*r;g^*})}$$

$$\leq \frac{\mathrm{P}}{S}\sum_{s,r}\sqrt{S^{-1}V(\tilde{X}^+_{its;k}\tilde{X}^+_{i^*t^*s;k^*})}\sqrt{S^{-1}V(\tilde{X}^+_{jmr;g}\tilde{X}^+_{j^*m^*r;g^*})} = o_p(1).$$

The first summation is simply the $S$ element-by-element variances in each spatial location; the equation shows that this is bound in probability by zero. The second summation consists of the $S$ variances plus the absolute value of the covariances. The second and third inequality hold due to A4. The last equality follows from $\lim_{S\to\infty}\mathrm{P}/S = 0$ along with A5 and A6. The asymptotic normality result, $\sqrt{S}(\hat{\beta}_* - \beta_*)\xrightarrow{d} N(0, M^{-1}\Gamma M^{-1})$, follows from White (1984, Theorem 4.25).

Together A1-6 are a linear version of the assumptions for asymptotic normality in Pinkse, Slade and Shen (2006) with some noticeable simplifications caused by (among other things) the linear form in (2) and the closed-form of the estimator. Therefore, $\hat{\beta}_*$ is asymptotically normal with $\sqrt{S}$ convergence rate. The blocking strategy also accommodates asymptotic normality as $N\to\infty$, $NS\to\infty$, $TS\to\infty$ or as $NTS\to\infty$, so the aforementioned asymptotics could be adjusted to accommodate a variety of convergence rates, as long as the dependences in the alternative dimensions can be adapted to those implied by A3 and A4. In cases where the time dimension grows, the usual temporal mixing conditions (e.g., White 1984, Definition 3.42) are essentially replaced with the spatial mixing conditions. Consistency of $\hat{\beta}_*$ is implied by the conditions for asymptotic normality (White, 1984, Theorem 2.28). Finally, for inference let the residual be $\hat{v}^+_s = Y^+_s - \tilde{X}^+_s\hat{\beta}_*$. Then a robust variance estimate in the spirit of White (1980) and Arellano (1987) is:

$$V(\hat{\beta}_*) = \left(\sum_{s=1}^{S} \tilde{X}_s^{+'} \tilde{X}_s^{+}\right)^{-1} \left(\sum_{s=1}^{S} \tilde{X}_s^{+'} \hat{v}_s^{+} \hat{v}_s^{+'} \tilde{X}_s^{+}\right) \left(\sum_{s=1}^{S} \tilde{X}_s^{+'} \tilde{X}_s^{+}\right)^{-1}. \quad (3)$$

We now mention some practical issues related to asymptotics along the spatial dimension. The increasing-domain asymptotics are problematic in a practical sense. To see this, we only need realize that as the lattice gets larger, there will not be enough time in period $t$ to move production to all (or a large number) of the spatial locations; there is just not enough time to travel the large distances. For example, if we are discussing fishing vessels, and the unit of $t$ is one week, and one vessel can fish a maximum of 25 different locations in one week, then expanding the number of locations above 25 for asymptotics is impractical. To remedy this we could expand the unit of observation for $t$ by aggregating across $t$. To continue the example, suppose we aggregated 52 weeks of weekly data into 12 months of monthly data, then over the course of a month a vessel may be able to visit four times as many spatial locations, so we could expand the maximal number of locations to 100. Now, we effectively have $S \to \infty$ while $T \to 0$ as our asymptotic argument. However, as $T \to 0$, we still have a problem, since, $\alpha_{it} \to \alpha_i$, and the model will be misspecified, as technical efficiency is no longer time-varying. We can also think of this as a violation of the fact that over large units of time, it is not practical to think of technical efficiency as being time-invariant. The infill asymptotics approach is less problematic, but there are still practical difficulties associated with it. If we divide the lattice into smaller and smaller spatial areas while keeping its total area fixed, then the lattice becomes a spatial continuum of fixed size in $s$. Unfortunately production data are inherently discrete in $s$, so increasing $S$ will eventually cause the production data at each location at be unmeasurable (in a discrete sense).

One could also envision some combination of these two asymptotic extremes. The spatial lattice is expanding while the spatial resolution is simultaneously increasing. This may provide some empirical benefits. For a particular data set, we may have large enough $S$ to appeal to asymptotics, where the lattice is not too big, so as to force $T$ to be too small to preclude time-varying technical efficiency, *and* where the spatial resolution is not too fine, so as to preclude data collection in each spatial location or to cause inputs to be fixed over space. Ultimately, adjusting the data through aggregation, disaggregation, or spatial normalization are empirical decisions that must balance time, space, and the dimensionality of the $\alpha_{it}$. Of course any aggregation along the time or spatial dimensions, will induce heteroskedasticity in the aggregate errors, so robust estimation is required.

**4. Application to Bearing Sea Flat Fisheries**

To illustrate our method, we use data on flatfish catch for 12 bottom trawl vessels within the Bering Sea from 2002 through 2004.[9] The data come from three sources. The spatial dimension of the data set is defined by the Alaska Department of Fish and Games (ADF&G) spatial locations, which partition the Bering Sea into grids that are one-half degree latitude by one-degree longitude in dimension. This produces approximately 95 spatial locations in the sea, but the average vessel only visits about 44 of these in a given year.[10] Production data (catch), $Y_{its} = \ln Catch_{its}$, is obtained from the National Marine Fisheries Service (NMFS) "observer program," which requires all vessels longer than 125 feet to have an observer onboard to record catch size, composition, and geographic position. On any given fishing trip, not all the catch is

[9] The primary target species harvested within the flatfish fishery is yellow-fin sole, however several additional species are harvested. These species are flathead sole, rock sole, rex sole, Greenland turbot, etc.. To simplify our example we aggregate all flatfish species captured into a single output. Initially we had 5 years of data, from 2000-2005. However, poolability tests indicated a structural break between 2001 and 2002, so we focus our example on the most recent portion of the data

[10] The 12 vessels were selected using a spatial site filter, requiring a vessel to visit at least 32 spatial locations within each year of the data set. Therefore, our analysis is only for the most mobile vessels in the fleet.

recorded, because observers take periodic breaks for sleep and hygiene. However, if we can assume that unobserved catch is random, our estimates should remain unbiased. Weekly catch data for the twelve vessels were aggregated to annual data, resulting in 1590 observations. That is, 12 vessels, over 3 years, each visiting on average a little over 44 spatial locations per year.

A quick experiment demonstrates that there is considerable variability in $\ln Catch_{its}$ over the spatial dimension. Different aggregation schemes reveal that: the average catch for each of the 12 vessels was 796.7 tons of fish with a standard deviation of 134.7, the average catch in each of the three years was 3,186.7 tons of fish with a standard deviation of 212.9 tons of fish, while the average catch in each of the 95 spatial locations was 100.6 tons of fish with a standard deviation of 83.6 fish. The spatial dimension of the data possesses the highest coefficient of variation (83.1%), so the spatial panel specification of equation 1 is well-justified. Figures 2, 3, and 4 show aggregate catch densities for 2002, 2003, and 2004, respectively. These are kernel smoothed surface plots of the 95 spatial locations with red areas indicating highest aggregate catch, yellow areas indicating medium catch, and blue areas indicating lowest catch for the 12 vessels in our sample. There is clearly a fair amount of spatial variability in output.

Observations in $X_{its}$ (also from the observer program data) are $\ln Hauls_{its}$ and $\ln Duration_{its}$, where $Hauls$ is the number of times the gear (a fish net) is deployed and $Duration$ is the total length of time that the gear is deployed. The spatially invariant inputs, $Z_{it}$ and $W_i$, are from the weekly production reports collected by NMFS as well as the United States Coast Guard vessel registry database, which records vessel characteristics. The $Z_{it}$ variable is $\ln Crew_{it}$ which is the logarithm of the "total number of crew members employed during the year divided by the number of weeks fished." The $W_i$ variable is $\ln NetTons_i$ which is the logarithm of net-tonnage of each vessel.[11] The data set is balanced across vessels and time but unbalanced across space.

Biomass densities, $b_{ts}$, are from the annual NMFS "biomass trawl survey," which biologist at the Alaska Fisheries Science Center (AFSC) use to calculate stock estimates. Annual stock assessment studies are conducted independently of the fishery (i.e., are not based on fishery output) and represent the best available estimates of the spatial distribution of the stock density. Figure 1 shows the spatial locations used in the analysis, which correspond to biomass survey points. In the case that $Catch$ is observed but $b_{ts}$ is not, then the mean biomass density within a given year is imputed. (This occurred in roughly 25% of the observations.)

We believe that our biomass data are exogenous, because a vessel captain's decision of "where to fish" is not based on this particular survey.[12] That is, catch incentives do not feedback into biomass through the harvest location decision. First, the annual stock assessment studies are conducted independently of the fishery (i.e., the stock assessment is not based on commercial catch). Also, Holland and Sutinen (2000) suggest that captains are "creatures of habit," tending to fish the same spatial pattern from year to year, regardless of survey data. Smith (2000) suggests that factors in the location decision are largely not observed by the analyst. Wilson (1990) suggests that fisheries have complex unobservable "informational networks" in which captains share location/catch information on a daily basis. Since stock measurements are taken annually, correlations between our biomass patterns and daily or hourly location decisions are negligible.[13]

---

[11] Data on vessel horsepower were available but not used due to high correlation with vessel net-tonnage.

[12] It may also be worth noting that if choice variables are endogenous by definition, then labor and capital are also endogenous, and the entire exercise of estimating a production function is not identified.

[13] There are bycatch biomass surveys conducted in this particular fisheries that *are* known to be used by captains in their location decisions. These surveys are based on vessel catch and are designed to help captains avoid bycatch species. However, this information is reported as a rate versus a raw amount.

The basic Cobb-Douglas harvesting function is:

$$\ln Catch_{its} = \alpha_{it} + b_{ts} \ln Hauls_{its} \beta_1 + b_{ts} \ln Duration_{its} \beta_2 + b_{ts} \ln Crew_{it} \gamma_1 + b_{ts} \ln NetTons \delta_1 + \varepsilon_{its}$$

Notice that each variable in interacted with biomass, $b_{ts}$, making them all space-varying (effectively).[14] The basic model was estimated and subjected to specifications test. Experimentation with interaction terms and a series of specifications tests led to the augmented Cobb-Douglas specification in Table 1, which includes the square of *Hauls* and an interaction between *NetTons* and *Crew*.[15] All tests were performed without accounting for aggregation heteroskedasticity. After the final specification was achieved, the standard errors were adjusted for heteroskedasticity. The t-statistics in Table 1 are based on equation 3. Since we know *ex ante* the form of the aggregation, we could correct our standard errors based on this structure, however White's correction is robust.

The results in Table 1 imply that the relationship between the number of hauls and production is nonlinear, and that crew size and vessel size (*NetTons*) are the only effective inputs to production insofar as they are appropriately mixed. That is, the positive coefficient of 0.0458 on *Crew*NetTons* implies that large vessels must have large crews and large crews must work on large vessels to be effective. Even though the coefficients on *Crew* and *NetTons* are negative, their elasticities are positive once we account for biomass and their interaction.

Elasticity estimates are contained in Table 2, and are transformed by average biomass over $t$ and $s$, $\bar{b} = 0.6937$. For example, the marginal product of *Crew* is:

$$\varepsilon_{Crew} = \frac{\partial \ln Y_{its}}{\partial \ln Crew_{its}} = \bar{b}\left[-0.7403_1 + 0.1150 \cdot \overline{\ln NetTons}\right],$$

[14] We experimented with a translog production function, but it was rejected by specifications tests.
[15] Some of the less parsimonious specifications had problems with highly collinear interactions. In cases where correlations exceeded 0.975, some interactions were eliminated from the specification.

where $\overline{\ln NetTons}$ is the average over $i$. The results imply that *Hauls* and *Duration* contribute more on the margin then any of the other inputs (0.4298 and 0.1934, respectively). *NetTons* provides the least (0.0501). However, all elasticities are positive, so our production model does not violate any of the traditional production theory assumptions. These results make sense. The act of deploying the nets (*Hauls*) and dragging the nets (*Duration*) is the most important input to harvesting fish. (Clearly, if this doesn't happen there will be zero output!) The next most important productive input to harvesting fish is crew size (elasticity of 0.0697); crews deploy and retrieve the nets. The size of the vessel is only important for speed and catch storage capacity, which are meaningless without a good crew and efficient deployment of the nets. Finally, returns to scale for the 12 vessels are 0.7430. The decreasing returns to scale may be from eliminating smaller vessels (below 125 feet), if these vessels exhibit constant or increasing returns.

Next, we estimated 36 different $\alpha_{it}$'s corresponding with 12 vessels and 3 years of data. These are in Table 3. The conditional variance matrix is,

$$V_1(\hat{\alpha}) = (G'G)^{-1}G'X_*^+V(\hat{\beta}_*)X_*^{+'}G(G'G)^{-1} + (G'G)^{-1}G'\hat{\Omega}G(G'G)^{-1},$$

where $\hat{\alpha} = [\alpha_{11} \quad \ldots \quad \alpha_{NT}]'$ and $G$ is the $(S \times NT)$ block diagonal matrix with typical diagonal $(S_{it} \times 1)$ block equal to $\iota_{S_{it}}$, an $S_{it} \times 1$ vector of ones. Based on this structure, we calculated the standard error of the difference in the individual effects between 2002 and 2004 (Table 3, column 5) and test statistics (column 6). We conclude that the only significant changes in effects between 2002 and 2004 are for vessels 1, 4, and 7.

Next, we are interested in performing simultaneous inference on the within year differences across vessels $u_{it}^* = \max_j \alpha_{jt} - \alpha_{it} > 0 \;\; \forall \;\; i = 1,...,12$. In other words, we want to know to what extent the vessel with the largest individual effect in the population (relative technical inefficiency equal to zero) dominates the other vessels in a given year, *simultaneously*. This test is indirectly performed using ranking and selection methods described in Horrace and Schmidt (2000). That is, assuming that the $\alpha_{it}$ are normal (or asymptotically so), we endeavor to determine a subset of the 12 vessels that contains the least inefficient vessel with probability 0.95. To do so, we simulate upper 95% percentage points for an 11-dimnesional multivariate normal distribution with means $\hat{\alpha}_{kt} - \hat{\alpha}_{it}$ for all $i \neq k$ and a general covariance structure, based on a linear transformations of $V_1(\hat{\alpha})$ for each year.[16] We do this for each $k$, producing 12 simulated critical values (percentage points), $z_{k,t}^{0.95}$ for $k = 1,...,12$ in each year $t$ = 1, 2, 3. The critical values are used to construct the three subsets:

$$\zeta_t = \{k\text{: } U_{it}^k \geq 0 \quad \text{for} \quad i \neq k \},$$

$$U_{it}^k = \hat{\alpha}_{kt} - \hat{\alpha}_{it} + z_{k,t}^{0.95}\{V(\hat{\alpha}_{kt}) + V(\hat{\alpha}_{it}) - 2Cov(\hat{\alpha}_{kt}, \hat{\alpha}_{it})\}^{1/2}, \quad t = 1, 2, 3.$$

The $U_{it}^k$ are 95% upper bounds on $\hat{\alpha}_{kt} - \hat{\alpha}_{it}$ for all $i \neq k$ for each $k$ in each year. A vessel $k$ belongs in $\zeta_t$ if it has all positive 95% upper bounds in year $t$. Then, $\zeta_t$ contain the indices of the vessels with the largest individual effect (smallest inefficiency) with probability 95% in each year. That is, let the index of the vessel with the (unknown) largest individual effect in year $t$ be $i_t^*$. Then,

$$\Pr\{i_t^* \in \zeta_t\} \geq 0.95, \quad t = 1, 2, 3.$$

That is, the index of the vessel with the largest individual effect (smallest inefficiency) is contained in $\zeta_t$ with probability at least 95%. The critical values, $z_{k,t}^{0.95}$, for each vessel $k$ for

[16] The linear transformation corresponds to the variance of the linear transformation of $\hat{\alpha}$ to $\hat{\alpha}_{kt} - \hat{\alpha}_{it}$, $k \neq i$.

each year $t$ are in the last three columns of Table 3 and are simulated using the algorithm in Horrace (1998) but for a general covariance structure on $\alpha_{it}$ and 100,000 simulation draws.

Based on these critical values, the 95% subsets of efficient vessels in each year are.

$\zeta_{2002}$ = {1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12},

$\zeta_{2003}$ = {3, 4, 5, 6, 7, 8, 9, 10, 11, 12},

$\zeta_{2004}$ = {3, 4, 5, 6, 8, 9, 10, 11, 12}.

The cardinality of the subsets is decreasing in time. As we move from 2002 to 2003, vessel 1 drops out of contention for the least inefficient boat, and as we move from 2003 to 2004, vessel 7 drops out of contention for the least inefficient boat. The drops correspond to the significant declines in $\alpha_{1t}$ and $\alpha_{7t}$ as we move from 2002 to 2004 in Table 3. For some reason these vessels had inefficiency increases over the period. Since it appears in none of the subsets, vessel 2 is never in contention for being least inefficient (most efficient). Notice that all the values for $\alpha_{2t}$ are lowest in each column in Table 3.

We now transform differences in the individual effects into efficiency estimates. Table 4 contains the time-varying technical efficiency measures for each vessel, based on the two methods described in section 2. The first column contains the vessel identifier. The second, third, and fourth columns contain the technical efficiency estimates, $T\hat{E}_{it} = \exp\{-\hat{u}_{it}\}$, which are relative to all vessels in all years. For example, the most efficient performance was vessel 4 in 2003. It's technical efficiency is normalized to 1.0. Remember, vessel 4 had a significant increase in its individual effect as we moved from 2002 to 2004. This is reflected in its efficiency scores of 0.4931, 1.0, and 0.8359 over the period. These efficiency scores are relative to its own performance in 2003. The alternative efficiency estimates $T\hat{E}^*_{it} = \exp\{-\hat{u}^*_{it}\}$ are in columns five, six, and seven of Table 4. These are within-year performance estimates. The most efficient vessels were 5, 4, and 11 in 2002, 2003, and 2004, respectively. The reader is reminded, however, that the t-tests in Table 3 suggest that efficiency difference across columns 2 through 7 are statistically insignificant in general. Also, most of the differences across rows of Table 4 are insignificant; this is reflected in the high cardinality of the three subsets, $\zeta_t$, that we saw earlier.

**5. Conclusions**

This research makes direct contributions to the panel data econometrics literature, the stochastic frontier literature, and the spatial econometrics literature. Highly-mobile technologies represent a very clean extension to the usual panel data results and add a degree of flexibility to asymptotic arguments and robust inference on model parameters. Our results are also meaningful for the stochastic frontier literature where estimation of time-varying individual effects is important. For a mobile technology we show that these parameters can be estimated without parametric assumptions and that large sample inference can be performed without an incidental parameters problem (even though we couldn't illustrate this in our example).

Our contribution to the spatial econometrics literature is clear. Our asymptotics results are a unique application of the spatial central limit theorems of Pinkse, Shen and Slade (2007) and Pinkse, Slade and Shen (2006). However, the results have implications for the estimation of spatial weighting matrices. It would be interesting to use the panel structure to estimate a spatial weighting matrix and compare it to the usual spatial weight matrix based on physical distance (e.g., Kelijian and Prucha, 1999 and 2001). Finally, our results may inform the location choice literature. For example, there are growing literatures on location choice in fisheries (e.g., Hick and Schnier, 2006; Smith, 2000; Smith and Wilen, 2003), agglomeration economies (e.g., Lovely, Rosenthal and Sharma, 2005), and migration (e.g., Dahl, 2002), that may benefit from the discussions herein.

Two weakness of the results are that resource stocks must be exogenous and that the individual effects cannot be space-varying. In the case that stocks are endogenous through the location decision, then appropriate instruments for stocks are necessary. In the case of U.S. fisheries, over the last few years, there have been important policy changes that have impacted the behavior of fishing vessels. Perhaps the timing of these exogenous policy changes, could be used as instruments. In fact, there are certain weekly or daily stock measures that are known to be used by vessel captains in their search for target fish species. Exploring policy changes as instruments for these stocks would be interesting. In the case where individual effects vary over both time and space our results do not apply, but an extension to the results of Ahn, Lee, and Schmidt (2004) would identify the model in a GMM framework. Also, the model could be identified with 'within' estimation if the individual effects where *time-invariant* but *space-varying*. In this case, interaction with resource stocks would be unnecessary, and the usual demeaning along the time dimension would produce the usual panel results. All these weaknesses will be addressed by the authors in subsequent research.

**References**

Ahn, S.C., Lee, Y.H., and P. Schmidt, 2001, GMM Estimation of Linear Panel Data Models with Time-Varying Individual Effects, *Journal of Econometrics* 101, 219-255.

Ahn, S.C., Lee, Y.H., and P. Schmidt, 2004, Panel data models with multiple time-varying individual effects, unpublished, department of economics, Hansung University.

Ahn, S.C., Lee, Y.H., and P. Schmidt, 2005, Panel data models with multiple time-varying individual effects: application to a stochastic frontier model, Unpublished, Hansung University.

Arellano, M., 1987. Computing robust standard errors for within-groups estimators, *Oxford Bulletin of Economics and Statistics*, 49, 431-34.

Battese G.E. and T.J. Coelli. 1992. Frontier Production Functions, Technical Efficiency, ad Panel Data. *Journal of Productivity Analysis* 3:153-169.

Bernstein, S., 1927. Sur l'extension du theoreme du calcul des probabilites aux sommes de quantites dependants, Mathematische Annalen 97, 1–59.

Chamberlain, G., 1984. Panel Data in Griliches, Z. and Intrilligator, M.D. (Eds.) Handbook of Econometrics, Vol. 2,

Chamberlain, G., 1992. Effciency bounds for semiparametric regression. Econometrica 60, 567-596.

Conley T. G. 1999. GMM Estimation with Cross Sectional Dependence, *Journal of Econometrics*, Vol. 92, 1-45.

Conley, T. G. and B. Dupor 2003. A Spatial Analysis of Sectoral Complementarity, *Journal of Political Economy*, Vol. 111 No. 2, 311-352.

Conley T. G. and E. A. Ligon 2002. Economic Distance, Spillovers, and Cross Country Comparisons, *Journal of Economic Growth*, Vol. 7, 157-187.

Conley T. G. and G. Topa 2002. Socio-economic Distance and Spatial Patterns in Unemployment, *Journal of Applied Econometrics,* Vol 17, Issue 4, 303-327.

Cornwell, C., Schmidt, P., and R. Sickles, 1990, Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels, *Journal of Econometrics* 46, 185-200.

Cressie, N.A.C. 1993. *Statistics for Spatial Data,* Wiley Interscience, NYC.

Cuesta, R.A. 2000. A Production Model with Firm-Specific Temporal Variation in Technical Inefficiency: with Application to Spanish Dairy Farms, *Journal of Productivity Analysis* 13, 139-149.

Dahl, G.B. 2002. Mobility and the Return to Education: Testing a Roy Model with Multiple Markets, *Econometrica*, 70, 2367-2420.

Davies A. and K. Lahiri. 1995. A new framework for analyzing survey forecasts using three-dimensional panel data , *Journal of Econometric*s, 68, 205-227.

Eilat, Y. and L. Einav. 2004. Determinants of international tourism: a three-dimensional panel data analysis, *Applied Economics*, 36, 1315-1327.

Feng, Q. and W.C. Horrace. 2007. Fixed-effect estimation of technical efficiency with time invariant dummies, Economics Letters, 95, 247-252.

Fernandez, C., Koop, G., Steel, M.F.J., 2002. Multiple output production with undesirable outputs: an application to nitrogen surplus in agriculture. *J. of the American Statistical Association* 97, 432–442.

Hall, R.E., Mishkin, F.S., 1982. The sensitivity of consumption to transitory income: estimates from panel data on households. *Econometrica* 50, 461-481.

Hansen, C. 2005. Asymptotic properties of a robust variance matrix estimator for panel data when *T* is large. *Journal of Econometrics*. forthcoming.

Hicks, R. L. and K. E. Schnier. 2006. Dynamic Random Utility Modeling: Monte Carlo Analysis, *American Journal of Agricultural Economics* 88(4) 816-835.

Holtz-Eakin, D., Newey, W.K., Rosen, H., 1988. Estimating vector autoregressions with panel data.*Econometrica* 56, 1371}1395.

Horrace W,C. 1998. Tables of percentage points of the k-variate normal distribution for large values of *k*. *Communications in Statistics: Simulation & Computation*, 27, 823-831.

Horrace W.C., Schmidt P., 2000. Multiple comparisons with the best, with economic applications. *Journal of Applied Econometrics*, 15, 1-26.

Han, C., Orea, L., and P. Schmidt, 2005, Estimation of a Panel Data Model with Parametric Temporal Variation in Individual Effects, *Journal of Econometrics*, 126 (2), 241-267.

Hausman, J.A. and W.E. Taylor. “Panel Data and Unobservable Individual Effects.” *Econometrica* 61(November 1981):1377-98.

Holland, D.S. and J.G. Sutinen, 2000. Location choice in New England trawl fisheries, *Land Economics*, 76, 133-149.

Holland, D.S., and J.G. Sutinen. 1999. “An Empirical Model of Fleet Dynamics in New England Trawl Fisheries.” *Canadian Journal of Fisheries and Aquatic Science* 56, 253–64.

Holtz-Eakin, D., Newey, W.K., Rosen, H., 1988. Estimating vector autoregressions with panel data. *Econometrica* 56, 1371-1395.

Keane, M.P., Runkle, D.E., 1992. On the estimation of panel data models with serial correlation when instruments are not strictly exogenous. *Journal of Business and Economic Statistics* 10, 1-10.

Kelejian H. H. and I. R. Prucha, 1999. A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model, *International Economic Review*, Vol. 40, 509-533.

Kelejian H. H. and I. R. Prucha, 2001. On the Asymptotic Distribution of the Moran I Test Statistic with Applications, *Journal of Econometrics*, Vol. 104, 219-257.

Kezdi, G., 2003. Robust Standard Errors Estimation in Fixed-Effects Panel Models, Unpublished, Budapest University of Economics.

Kiefer, N.M., 1980. A time series-cross section model with fixed effects with an intertemporal factor structure. Unpublished manuscript, Cornell University.

Kim, Y., Schmidt, P., 2000. A review and empirical comparison of Bayesian and classical approaches to inference on efficiency levels in stochastic frontier models with panel data. *Journal of Productivity Analysis* 14, 91–118.

Kneip, A., R.C. Sickles, and W. Song  2005. A New Panel Data Treatment for Heterogeneity in Time Trends, Unpublished, Department of Economics Rice University.

Koop, G., Osiewalski, J., Steel, M.F.J., 1997. Bayesian efficiency analysis through individual effects: hospital cost frontiers. *Journal of Econometrics*, 77–105.

Kumbhakar, S.C., 1990. Production frontiers, panel data, and time-varying technical inefficiency. *Journal of Econometrics* 46, 201–211.

Lee, Y.H., 2005, A Stochastic Production Frontier Model with Group-Specific Temporal Variation in Technical Efficiency, *European Journal of Operational Research*, forthcoming.

Lee, Y.H., and P. Schmidt, 1993, A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency, In: Fried, H., Lovell, C.A.K., Schmidt, P. (Eds.) The

Measurement of Productiove Efficiency: Techniques and Applications, Oxford University Press, Oxford.

Lovely, M.E., S.S. Rosenthal, and S. Sharma, 2005. Information, Agglomeration and the Headquarters of U.S. Exporters," *Regional Science and Urban Economics*, 35, 167-191.

MaCurdy, T.E., 1981. An empirical model of labor supply in a life-cycle setting. Journal of Political Economy 89, 1059-1086.

Mundlak, Y., 1978. On the pooling of time series and cross section data. Econometrica 46, 69-85.

Newey, W.K., 1985. Generalized method of moments specification testing. Journal of Econometrics 29, 229-256.

Park, B. R. Sickles and L. Simar, 1998, Stochastic Panel Frontiers: A Semiparametric Approach, *Journal of Econometrics*, 84, 273-301.

Parsley, D.C. and S-J Wei. 2001. Explaining the border effect: the role of exchange rate variability, shipping costs, and geography. *Journal of International Economics*, 55, 87-105.

Pinkse, J., M. Slade and L. Shen, (2006). Dynamic spatial discrete choice using one step GMM: an application to mine operating decision. *Spatial Economic Analysis*, 1, 53-99.

Pinkse, J., L. Shen, and M. Slade 2007. A central limit theorem for endogenous locations and complex spatial interaction.. *Journal of Econometrics*, 140, 215–225.

Sanchirico, J. and J.E. Wilen. 1999. Bioeconomics of spatial exploitation in a patchy resource environment. *Journal of Environmental Economics and Management* 37:129-150.

Sanchirico, J. and J.E. Wilen. 2005. Optimal Spatial Management of Renewable Resources: Matching Policy Scope to Ecosystem Scale. *Journal of Environmental Economics and Management* 50: 23-46.

Schmidt, P. and Sickles, R. 1884. Production Frontiers and Panel Data, *Journal of Business and Economics Statistics*, 2, 367-374.

Schaefer, M.B. 1957 Some considerations of population dynamics and economics in relation to the management of marine fisheries. *Journal of the Fisheries Research Board of Canada* 14, 669-81.

Shapiro, M.D., 1984. The permanent income hypothesis and the real interest rate: some evidence from panel data. Economics Letters 14, 93-100.

Shephard, R.W. 1970. *Theory of Cost and Production Functions*, Princeton University Press.

Smith, M.D. 2000. Spatial search and fishing location choice: Methodological challenges of empirical modeling. *American Journal of Agricultural Economics* 82: 1198-1206.

Smith, M.D. and J.E. Wilen. 2003. Economic impacts of marine reserves: the importance of spatial behavior. *Journal of Environmental Economics and Management* 46: 183-206.

Tsionas, E.G., 2002. Stochastic frontier models with random coefficients. *Journal of Applied Econometrics* 17, 127–14

White, H. 1980. A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. Econometrica 48, 817-38.

White, H. 1984. *Asymptotic Theory for Econometricians*. Academic Press. New York.

Wilson, J.A., 1990. Fishing for Knowledge. Land Economics, 66, 12–29.

Table 1: Model Parameter Estimates

| Variable | Coefficient (t-statistic) |
|---|---|
| $b_{ts} \ln Hauls_{its}$ | 0.3758* (1.80) |
| $\frac{1}{2} b_{ts} (\ln Hauls_{its})^2$ | 0.1257 (1.52) |
| $b_{ts} \ln Duration_{its}$ | 0.2788** (1.96) |
| $b_{ts} \ln Crew_{it}$ | -0.7403** (-2.79) |
| $b_{ts} Crew_{it} * NetTons_i$ | 0.1150** (2.63) |
| $b_{ts} NetTons_i$ | -0.4612* (-1.79) |

*indicates significance at the 90% level.
**indicates significance at the 95% level.
t-statistics are heteroskedasticity robust.
$(\ln Hauls)^2$ was significant *before* standard error correction, so it is included in the specification.

Table 2: Elasticities and Returns to Scale

| $\varepsilon_{Hauls}$ | $\varepsilon_{Duration}$ | $\varepsilon_{Crew}$ | $\varepsilon_{Net-tons}$ | Returns-to-scale (RTS) |
|---|---|---|---|---|
| 0.4298 | 0.1934 | 0.0697 | 0.0501 | 0.7430 |

Table 3. Time-Varying Individual Effects and Critical values for Subset Selections

| Vessel | 2002 $\hat{\alpha}_{i1}$ | 2003 $\hat{\alpha}_{i2}$ | 2004 $\hat{\alpha}_{i3}$ | Stand. Error $(\hat{\alpha}_{i3}-\hat{\alpha}_{i1})$ | t-statistic | 2002 $z^{k,0.95}_{V(\hat{\alpha}),1}$ | 2003 $z^{k,0.95}_{V(\hat{\alpha}),2}$ | 2004 $z^{k,0.95}_{V(\hat{\alpha}),3}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 5.8477 | 5.5425 | 5.3933 | 0.2466 | -1.84* | 2.7193 | 2.7365 | 2.7510 |
| 2 | 3.4038 | 3.1069 | 3.6001 | 0.2719 | 0.72 | 2.7108 | 2.7383 | 2.6808 |
| 3 | 6.2102 | 6.3210 | 6.2651 | 0.2396 | 0.23 | 2.7569 | 2.7425 | 2.7200 |
| 4 | 5.8973 | 6.6044 | 6.4251 | 0.2186 | 2.41** | 2.7587 | 2.7422 | 2.7563 |
| 5 | 6.4230 | 6.1717 | 6.4193 | 0.2272 | -0.02 | 2.7478 | 2.7465 | 2.7617 |
| 6 | 6.0221 | 6.0868 | 6.1238 | 0.2303 | 0.44 | 2.7518 | 2.7644 | 2.7493 |
| 7 | 6.3429 | 6.2018 | 5.7982 | 0.2310 | -2.36** | 2.7437 | 2.7488 | 2.7312 |
| 8 | 6.2138 | 6.0815 | 6.1089 | 0.2437 | -0.43 | 2.7221 | 2.7049 | 2.7436 |
| 9 | 6.3282 | 6.2119 | 6.1578 | 0.2187 | -0.78 | 2.7594 | 2.7694 | 2.7621 |
| 10 | 6.2710 | 6.3304 | 6.0381 | 0.2378 | -0.98 | 2.7510 | 2.7431 | 2.7460 |
| 11 | 6.3773 | 6.4074 | 6.5328 | 0.2346 | 0.66 | 2.7308 | 2.7414 | 2.7568 |
| 12 | 6.2470 | 6.2928 | 6.4846 | 0.2368 | 1.00 | 2.7453 | 2.7521 | 2.7532 |

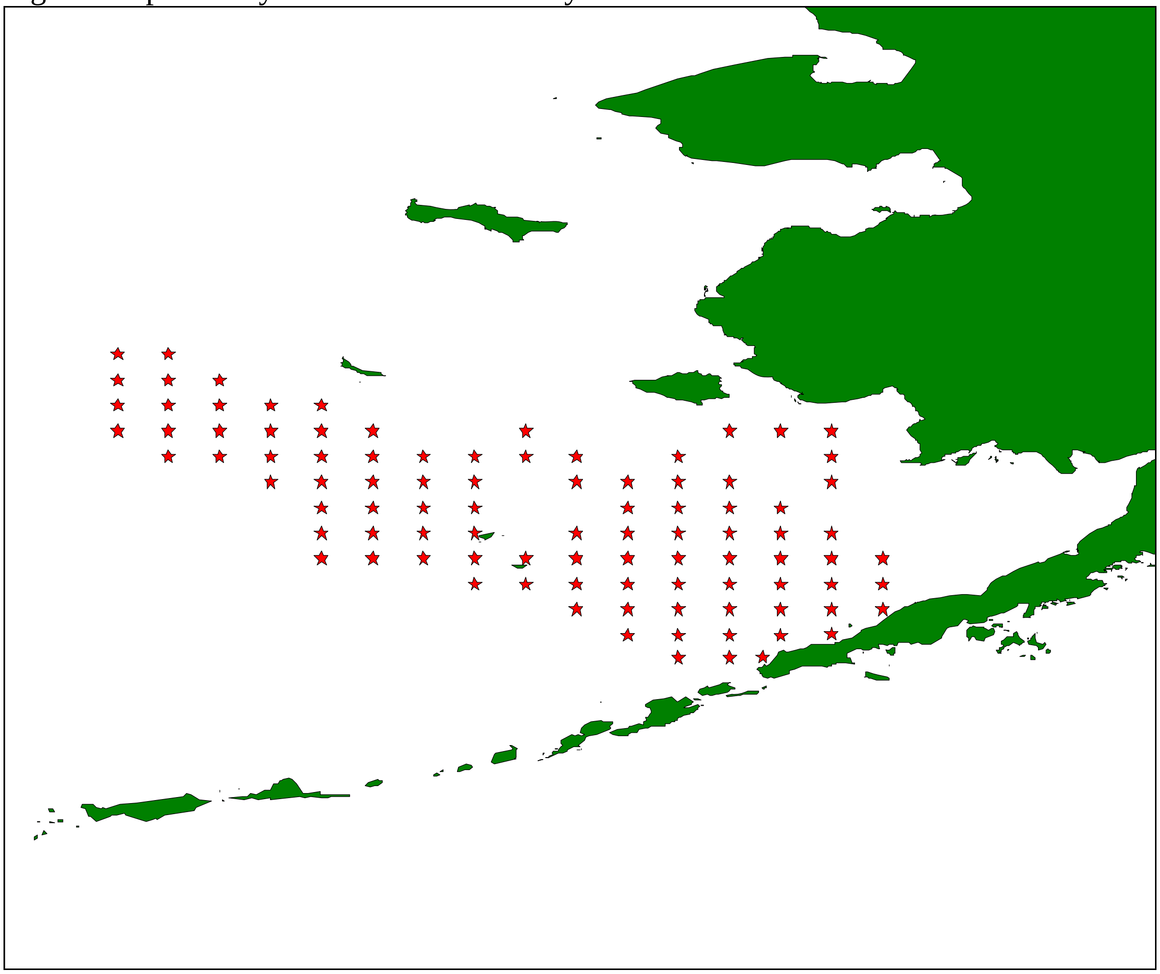
t-statistic for $H_0 : \hat{\alpha}_{i3} = \hat{\alpha}_{i1}$;
*indicates significance at the 90% level.
**indicates significance at the 95% level.
Critical values simulated for a general covariance structure per Horrace and Schmidt (2000).
$\zeta_{2002}$ = {1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}.
$\zeta_{2003}$ = {3, 4, 5, 6, 7, 8, 9, 10, 11, 12}.
$\zeta_{2004}$ = {3, 4, 5, 6, 8, 9, 10, 11,12}.

Table 4: Technical Efficiency and Number of Spatial Sites Visited Each Year by Vessel

| Vessel | 2002 $\hat{TE}_{it}$ | 2003 $\hat{TE}_{it}$ | 2004 $\hat{TE}_{it}$ | 2002 $\hat{TE}^{*}_{i1}$ | 2003 $\hat{TE}^{*}_{i2}$ | 2004 $\hat{TE}^{*}_{i3}$ |
|---|---|---|---|---|---|---|
| 1 | 0.4692 | 0.3458 | 0.2979 | 0.5625 | 0.3458 | 0.3200 |
| 2 | 0.0407 | 0.0303 | 0.0496 | 0.0488 | 0.0303 | 0.0533 |
| 3 | 0.6742 | 0.7533 | 0.7122 | 0.8083 | 0.7533 | 0.7651 |
| 4 | 0.4931 | 1.0 | 0.8359 | 0.5911 | 1.0 | 0.8979 |
| 5 | 0.8341 | 0.6488 | 0.8311 | 1.0 | 0.6488 | 0.8927 |
| 6 | 0.5586 | 0.5960 | 0.6184 | 0.6697 | 0.5960 | 0.6643 |
| 7 | 0.7699 | 0.6686 | 0.4466 | 0.9230 | 0.6686 | 0.4797 |
| 8 | 0.6767 | 0.5928 | 0.6093 | 0.8112 | 0.5928 | 0.6545 |
| 9 | 0.7587 | 0.6754 | 0.6398 | 0.9095 | 0.6754 | 0.6873 |
| 10 | 0.7165 | 0.7603 | 0.5676 | 0.8590 | 0.7603 | 0.6098 |
| 11 | 0.7968 | 0.8212 | 0.9309 | 0.9553 | 0.8212 | 1.0 |
| 12 | 0.6995 | 0.7323 | 0.8871 | 0.8386 | 0.7323 | 0.9529 |

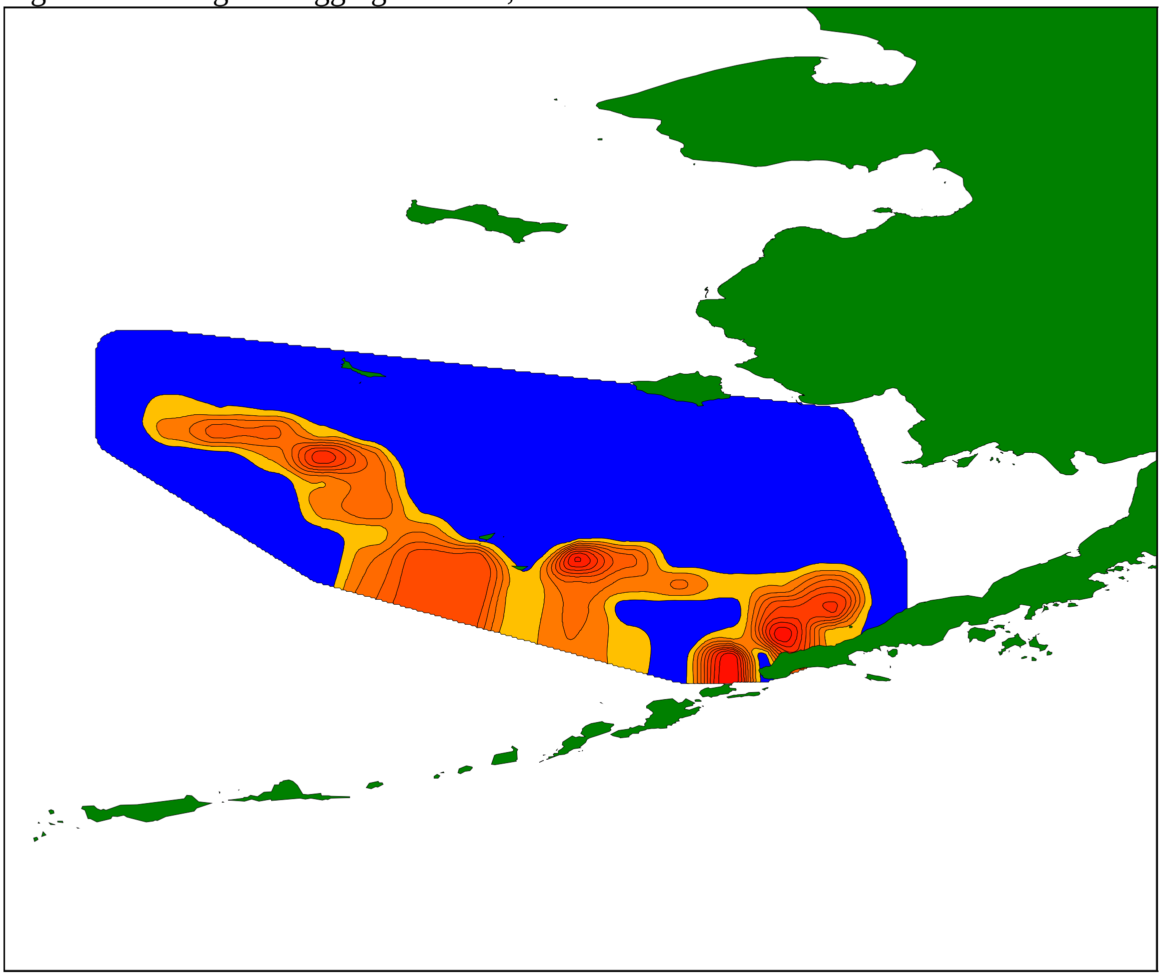
$\hat{TE}_{it}$ is relative technical efficiency across all vessels in all years
$\hat{TE}^{*}_{it}$ is relative technical efficiency across all vessels within year *t*.

Figure 1: Spatial Layout of the Trawl Survey Data



Each star indicates one of 95 spatial locations fished

Figure 2. Bearing Sea Aggregate Catch, 2002



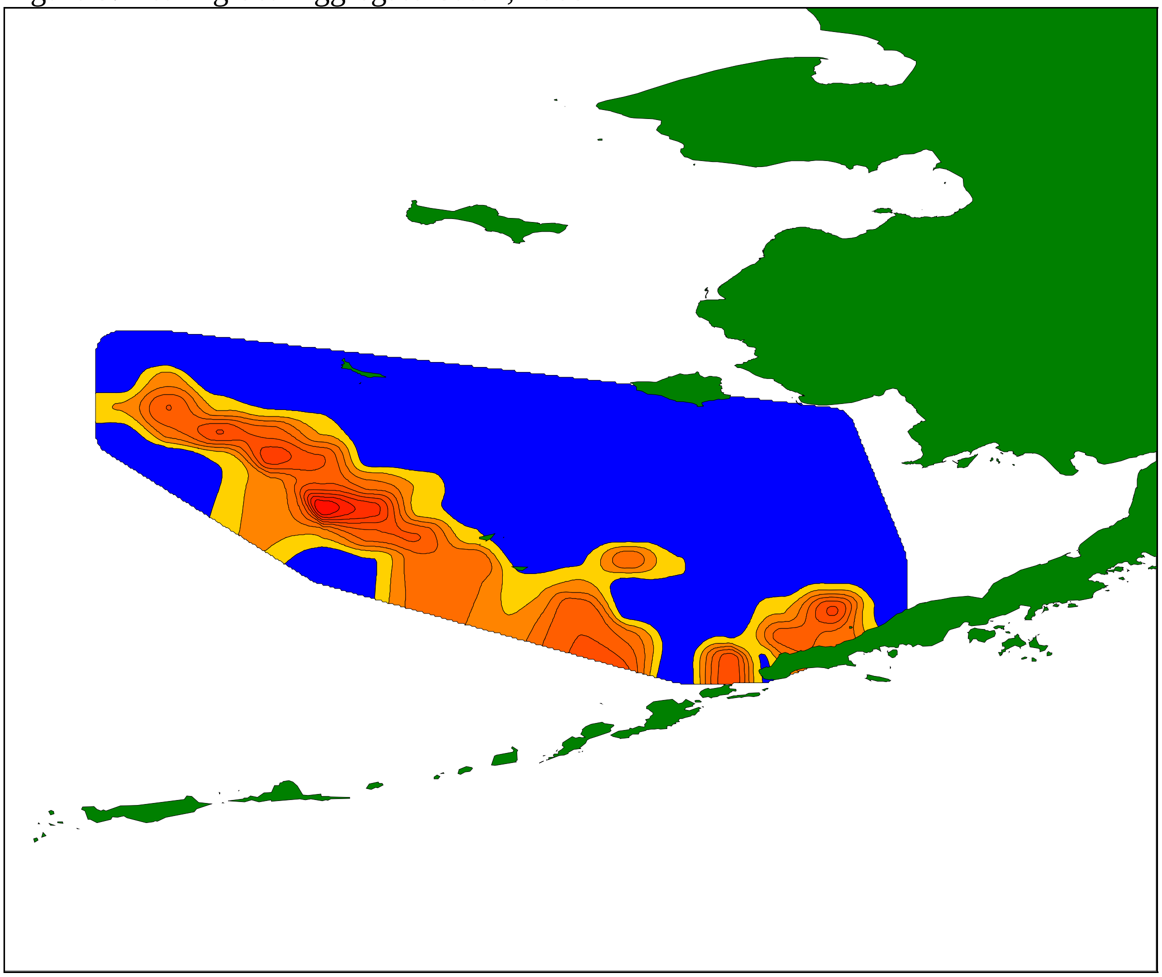
Kernel smooth of discrete location data.

= highest catch

= medium catch

= lowest catch

Figure 3. Bearing Sea Aggregate Catch, 2003



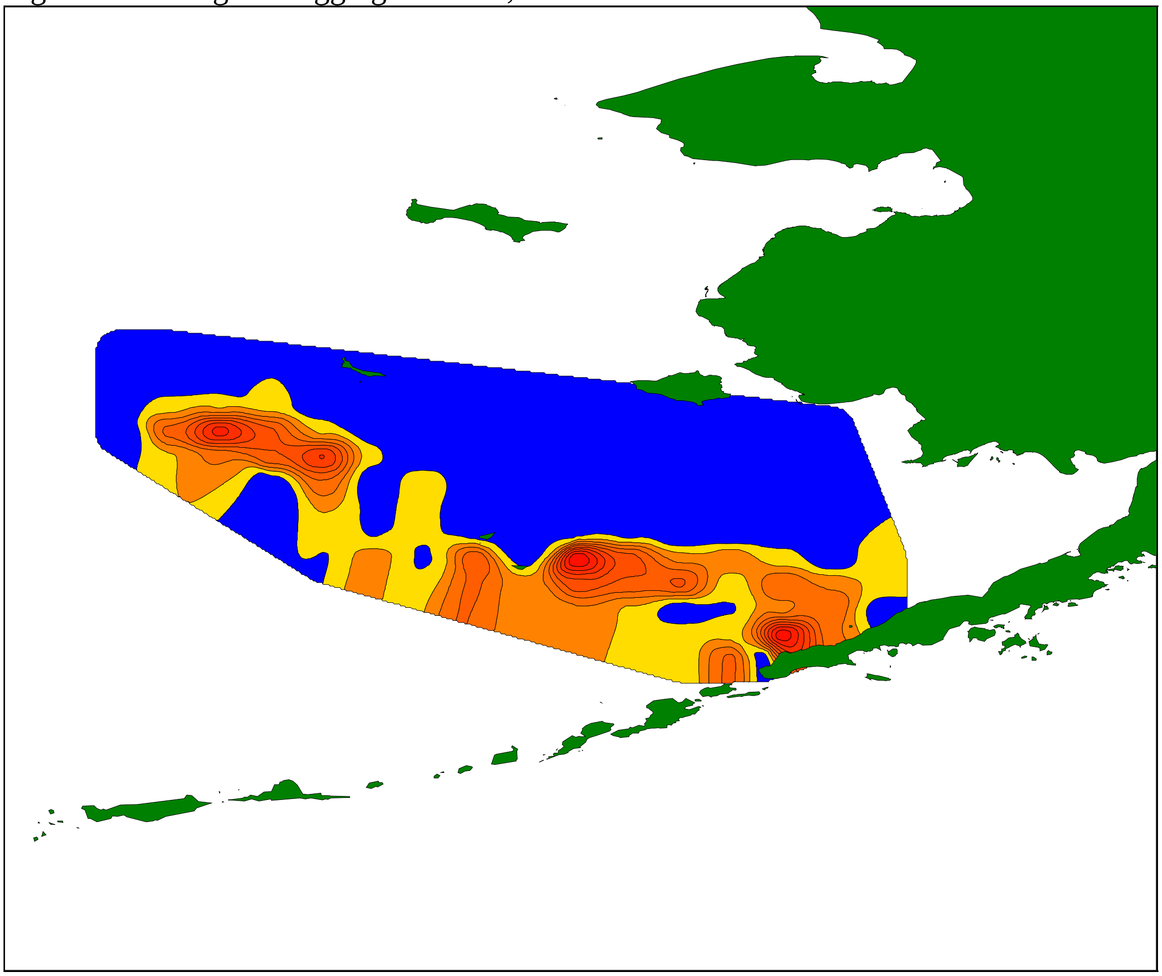
Kernel smooth of discrete location data.

= highest catch

= medium catch

= lowest catch

Figure 4. Bearing Sea Aggregate Catch, 2004



Kernel smooth of discrete location data.

= highest catch

= medium catch

= lowest catch